

# یک روش هوشمند جدید مبتنی بر شبکه‌های کپسولی برای طبقه‌بندی کشتی‌های دریایی

حامد جباری<sup>۱</sup>، حمید هوشمند<sup>۲</sup> و نوشین بیگدلی<sup>۳</sup>

## چکیده

شناسایی و طبقه‌بندی دقیق کشتی‌ها نقش مهمی در پایش هوشمند ترافیک دریایی، امنیت دریایی و مدیریت عملیات امداد و نجات دارد. با این حال، پیچیدگی محیط‌های دریایی، تغییرات جوی، شرایط نوری نامناسب و شباهت ظاهری کشتی‌ها، چالش‌های جدی برای سیستم‌های شناسایی ایجاد می‌کند. در این مقاله، یک روش هوشمند جدید مبتنی بر شبکه‌های کپسولی برای طبقه‌بندی تصاویر کشتی‌های دریایی ارائه شده است. شبکه‌های کپسولی با استفاده از کپسول‌ها به‌عنوان واحدهای پردازش اطلاعات، قادر به حفظ روابط فضایی بین ویژگی‌هاست و اطلاعات جامع‌تری از ساختار اشیاء ارائه می‌دهند. این ویژگی باعث افزایش صحت تصمیم‌گیری مدل می‌شود. برای ارزیابی عملکرد مدل پیشنهادی از مجموعه داده MARVEL با ۲۴ هزار تصویر در ۸ کلاس ناوهای هواپیمابر، باربرها، کشتی‌های جنگی، کشتی‌های کروز، کشتی‌های تندرو، کشتی‌های پلیس، کشتی‌های ماهیگیری و کشتی‌های لایروبی استفاده شد. در آزمایش‌های انجام شده، صحت کلی مدل در طبقه‌بندی تصاویر ۹۷/۶۱ درصد بود. در میان کلاس‌ها بیشترین صحت مربوط به کلاس کشتی‌های کروز با ۹۹ درصد و کمترین صحت مربوط به کلاس ناوهای هواپیمابر با ۹۵ درصد بود. همچنین نتایج نشان داد که مدل پیشنهادی در شرایط نوری متفاوت، فاصله‌های زیاد و وجود چندین شیء در تصویر عملکرد مناسبی دارد. مقایسه صحت، حساسیت و عملکرد زمانی مدل پیشنهادی با شبکه‌های عمیق شناخته‌شده مانند ResNet، GoogleNet و VGGNet نشان داد که مدل پیشنهادی در بسیاری از شرایط برتر عمل می‌کند. یافته‌ها نشان دادند که مدل ما ضمن عملکرد مناسب، می‌تواند به‌عنوان ابزاری قدرتمند و مؤثر برای طبقه‌بندی کشتی‌های دریایی در سامانه‌های نظارتی مورد استفاده قرار گیرد.

## کلید واژه‌ها

طبقه‌بندی، کشتی‌های دریایی، یادگیری عمیق، شبکه‌های کپسولی، شبکه‌های کانولوشنی.

## ۱- مقدمه

نقش زیست‌محیطی، اقیانوس‌ها به‌عنوان مسیرهای اصلی گردشگری و حمل‌ونقل بین‌المللی، از منظر تجاری، اقتصادی و امنیتی اهمیت بالایی دارند. با این حال، نظارت و حفاظت از گستره وسیع اقیانوس‌ها با محدودیت‌هایی همراه است. عواملی نظیر کندی جابه‌جایی کشتی‌های تحقیقاتی، شرایط ناپایدار جوی، و عمر محدود تجهیزات و حسگرها، چالش‌های جدی در این مسیر ایجاد می‌کنند [۲]. از طرفی، پایش گسترده‌ی ویدئوهای نظارتی دریایی مستلزم نظارت لحظه‌ای توسط خدمه یا نیروهای مدیریت ترافیک است؛ فرآیندی که به منابع انسانی و تجهیزات قابل توجهی تکیه دارد و در اثر خستگی یا خطای انسانی با چالش‌هایی مواجه می‌شود [۳]. مطالعات نشان می‌دهند که حدود ۸۰ تا ۹۰ درصد از حوادث دریایی ناشی از خطای خدمه است [۴]. در حال حاضر، توانایی‌های بینایی کامپیوتری برای کنترل

اقیانوس‌ها نقشی اساسی در تنظیم اقلیم زمین ایفا می‌کنند و به همین دلیل، نظارت بر وضعیت آنها و حفاظت از سلامت زیست‌محیطی‌شان از اهمیت ویژه‌ای برخوردار است. هم‌چنین، این پهنه‌های آبی گسترده، منابع طبیعی و معدنی گران‌بهایی را در خود نهفته دارند که صیانت از آنها را ضروری می‌سازد [۱]. افزون بر

این مقاله در اسفندماه ۱۴۰۳ دریافت شد؛ در اردیبهشت‌ماه ۱۴۰۴ بازنگری و در خردادماه پذیرفته گردید.

<sup>۱</sup> استادیار، گروه مهندسی برق کنترل، دانشگاه بین‌المللی امام خمینی (ره) قزوین  
رایانامه: [H\\_Jabbari@edu.ikiu.ac.ir](mailto:H_Jabbari@edu.ikiu.ac.ir)

<sup>۲</sup> استادیار، دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، سمنان  
رایانامه: [Hamidhooshmand85@gmail.com](mailto:Hamidhooshmand85@gmail.com)

<sup>۳</sup> استاد، گروه مهندسی برق کنترل، دانشگاه بین‌المللی امام خمینی (ره) قزوین  
رایانامه: [N.Bigdeli@eng.ikiu.ac.ir](mailto:N.Bigdeli@eng.ikiu.ac.ir)

نویسنده مسئول: حامد جباری

شباهت بالا میان کلاس‌ها مثل تصاویر اهداف دریایی، منجر به کاهش صحت طبقه‌بندی شود [۹]. در این پژوهش، برخلاف مطالعات پیشین از جمله [۱۰]، [۱۱]، [۱۲] و [۱۳] که عمدتاً از شبکه‌های کانولوشنی کلاسیک برای طبقه‌بندی کشتی‌های دریایی استفاده کرده‌اند، یک مدل نوآورانه مبتنی بر شبکه‌های کپسولی<sup>۵</sup> (CapsNet) طراحی شده است. نوآوری اصلی این پژوهش در به‌کارگیری شبکه‌های کپسولی برای حفظ ساختار فضایی بین اجزای کشتی در تصاویر دریایی و غلبه بر چالش‌های ناشی از پس‌زمینه‌های پیچیده، شرایط نوری متغیر، و شباهت زیاد بین کلاس‌های مختلف کشتی‌هاست. مدل پیشنهادی با بهره‌گیری از بردارهای ویژگی در کپسول‌ها و سازوکار مسیریابی پویا<sup>۱۴</sup> قادر است به طور دقیق‌تری ویژگی‌های مکانی و هندسی اهداف دریایی را استخراج و حفظ کند. این ویژگی باعث کاهش نرخ خطا در طبقه‌بندی کشتی‌های مشابه و افزایش صحت در شرایط جوی و نوری نامساعد شده است. هم‌چنین مدل پیشنهادی با بهره‌گیری از یک لایه ترکیبی چندمسیره و چندمقیاسی، امکان استخراج هم‌زمان ویژگی‌های محلی و سراسری تصاویر را فراهم می‌سازد. این لایه با استفاده از مسیرهای موازی شامل فیلترهایی با اندازه‌های متنوع طراحی شده است تا بازنمایی دقیق‌تری از تصویر ایجاد کرده و توان مدل را در شناسایی الگوهای متنوع موجود در تصاویر دریایی افزایش دهد. نتایج تجربی نیز نشان داد که روش ارائه‌شده در مقایسه با مدل‌های مرجع عملکرد بهتری داشت و صحت طبقه‌بندی اهداف دریایی را افزایش داد.

ساختار این مقاله به این شرح است: بخش دوم به مرور کارهای پیشین در زمینه طبقه‌بندی اهداف دریایی اختصاص دارد. در بخش سوم، روش پیشنهادی ارائه می‌شود. بخش چهارم نتایج به‌دست‌آمده را شرح می‌دهد و بخش پنجم به بحث و تحلیل این نتایج می‌پردازد. در نهایت، بخش ششم به جمع‌بندی یافته‌ها و نتیجه‌گیری مقاله اختصاص یافته است.

## ۲- کارهای پیشین

در سال‌های اخیر، روش‌ها و الگوریتم‌های متنوعی برای طبقه‌بندی کشتی‌ها مورد استفاده قرار گرفته است [۴، ۱۵]. جدول ۱ خلاصه‌ای از روش‌های موجود برای شناسایی و طبقه‌بندی کشتی‌های دریایی را همراه با مزایا و معایب آنها ارائه می‌دهد. به‌طور کلی، این روش‌ها را می‌توان به دو دسته اصلی تقسیم کرد: روش‌های مبتنی بر ویژگی و روش‌های مبتنی بر یادگیری عمیق. همان‌طور که در این جدول نشان داده شده است، روش‌های مبتنی بر ویژگی شامل مراحل پیش‌پردازش، تشخیص خط دریا-آسمان<sup>۶</sup> (SSL)، استخراج منطقه مطلوب<sup>۸</sup> (ROI) و طبقه‌بندی می‌شوند.

کشتی‌های سطحی خودران<sup>۱</sup> (ASV) و شناسایی خودکار اهداف<sup>۲</sup> (ATR) کافی نیست. از این رو، استفاده از فناوری‌های هوش مصنوعی برای تحلیل صحنه‌های دریایی و کمک به تصمیم‌گیری در کشتی‌های دارای خدمه و خودران ضروری است [۵، ۶]. ترکیب فناوری‌های هوشمند با داده‌ها و تصاویر عظیم اقیانوسی، امکان پایش دقیق‌تر و کارآمدتر را فراهم می‌کند. یکی از ابعاد کلیدی در پایش دریایی، شناسایی و طبقه‌بندی اهداف دریایی است؛ فرآیندی که در حوزه‌های نظامی و تجاری از اهمیت بسزایی برخوردار است [۲]. اهداف دریایی می‌توانند شامل انواع کشتی‌ها، زیردریایی‌ها، پهپادها، سامانه‌های پرتاب موشک و دیگر تجهیزات فعال در محیط‌های آبی باشند. با توجه به تنوع کشتی‌ها از نظر اندازه، شکل، و نوع کاربرد در زمینه‌هایی نظیر مهندسی دریایی، حمل‌ونقل، امنیت دریایی و پایش محیطی، طبقه‌بندی دقیق آنها ضرورتی انکارناپذیر است [۷]. تدوین استانداردها و مقررات مرتبط با هر نوع کشتی بر پایه‌ی طبقه‌بندی مناسب، نقشی مؤثر در مدیریت بهینه، افزایش ایمنی، و تسهیل حمل‌ونقل بار و مسافر در مسیرهای دریایی ایفا می‌کند. با این حال، طبقه‌بندی اهداف دریایی به‌ویژه کشتی‌ها با چالش‌های متعددی همراه است. مطالعات پیشین نشان داده‌اند که در شرایطی که تصاویر دریایی دارای پس‌زمینه‌های پیچیده هستند، شناسایی دقیق نوع کشتی با دشواری‌هایی همراه است [۴]. افزون بر این، حضور هم‌زمان اجسام دیگر یا کشتی‌های مجاور که موجب هم‌پوشانی در تصاویر می‌شوند، فرآیند تشخیص را پیچیده‌تر می‌سازد. شرایط نوری و جوی متغیر نیز می‌توانند منجر به شکل‌گیری اشیای مبهم یا نویز در تصاویر شوند که صحت شناسایی را کاهش می‌دهد. از دیگر چالش‌های مهم می‌توان به شباهت ظاهری بالا میان کلاس‌های مختلف کشتی‌ها و تنوع مدل‌ها در هر کلاس اشاره کرد؛ عواملی که احتمال بروز خطا در فرآیند طبقه‌بندی را به‌طور چشم‌گیری افزایش می‌دهند.

شبکه‌های عصبی عمیق<sup>۳</sup> (DNNs) در سال‌های اخیر به ابزارهای مؤثری برای تحلیل داده‌های پیچیده، به‌ویژه در حوزه پردازش تصویر و شناسایی الگوهای بصری، تبدیل شده‌اند [۸]. این شبکه‌ها با بهره‌گیری از چندین لایه یادگیری غیرخطی، قادرند ویژگی‌های سطح بالا را به‌صورت خودکار از داده‌های خام استخراج کنند و در مسائل طبقه‌بندی، شناسایی و تشخیص اشیاء عملکرد قابل‌توجهی ارائه دهند. با این حال، مدل‌های متداول شبکه‌های عمیق مانند شبکه‌های عصبی کانولوشنی<sup>۴</sup> (CNNs) در حفظ روابط فضایی میان اجزای تصویر محدودیت دارند. این مسئله می‌تواند به‌ویژه در طبقه‌بندی تصاویر با ساختار پیچیده و

<sup>۵</sup> Capsule Neural Networks

<sup>۶</sup> Dynamic Routing

<sup>۷</sup> Sky-Sea Line

<sup>۸</sup> Region of Interest

<sup>۱</sup> Autonomous Surface Vessel

<sup>۲</sup> Automatic Target Recognition

<sup>۳</sup> Deep Neural Networks

<sup>۴</sup> Convolutional Neural Networks

روش‌های مبتنی بر ویژگی‌های نیاز به تعداد زیادی نمونه آموزشی ندارند و از منابع محاسباتی نسبتاً کمتری استفاده می‌کنند. همچنین، این روش‌ها دارای پیاده‌سازی ساده هستند و قابلیت تفسیر بالایی دارند. با این حال، یافتن ویژگی‌های متمایز در موارد پیچیده چالش برانگیز است و گاهی نیاز به انتخاب ویژگی‌های دستی مناسب دارد. با پیشرفت شبکه‌های عصبی عمیق و افزایش توان مدل‌های یادگیری ماشین، روش‌های مبتنی بر یادگیری عمیق نیز برای طبقه‌بندی کشتی‌ها به کار گرفته شدند [۸]. این روش‌ها قادر به شناسایی الگوهای پیچیده و بهبود طبقه‌بندی از طریق استخراج خودکار ویژگی‌های ترکیبی و پیچیده از تصاویر هستند. الگوریتم‌های یادگیری عمیق می‌توانند هم روی تصاویر پیش‌پردازش شده و هم روی تصاویر خام اعمال شوند. با این حال، آموزش و استفاده از شبکه‌های عصبی عمیق معمولاً به حجم بالای داده‌های آموزشی و توان پردازشی بالا نیاز دارد. به‌طور کلی، روش‌های مبتنی بر یادگیری عمیق را می‌توان به سه دسته تقسیم کرد: یادگیری نظارت‌شده، یادگیری بدون نظارت، و یادگیری انتقالی. در یادگیری نظارت‌شده، داده‌های ورودی و خروجی (ویژگی-برچسب) مشخص هستند و هدف یافتن الگوها و روابط بین این داده‌ها است. بر اساس داده‌های موجود، یک مدل یادگیری ساخته می‌شود که توانایی پیش‌بینی یا طبقه‌بندی داده‌های جدید را دارد [۷]. شبکه‌های CNN، AlexNet، و ResNet نمونه‌هایی از این نوع یادگیری هستند. در یادگیری بدون نظارت، الگوریتم‌هایی مانند خودرمزگذارها برای کشف الگوها، اطلاعات مفید، و روابط پنهان در داده‌ها بدون نیاز به برچسب‌گذاری صریح به کار می‌روند [۳۴]. در یادگیری انتقالی، مدل‌ها و ویژگی‌های کسب‌شده در یک مسئله پایه (مسئله مبدأ)، برای بهبود یادگیری در مسئله هدف (مسئله مقصد) به کار گرفته می‌شوند. یادگیری انتقالی به دلیل مزایایی مانند کاهش نیاز به داده‌های آموزشی حجیم، کاهش زمان و هزینه آموزش، و بهبود عملکرد در مسائل با داده‌های محدود، جایگاه ویژه‌ای در یادگیری ماشین دارد [۲۶].

تصاویر دریافتی معمولاً حاوی نویز و اشیای ناخواسته هستند، بنابراین اولین مرحله در این روش‌ها، حذف نویز از تصاویر است. در پژوهش‌های پیشین، از روش‌هایی مانند تکنیک‌های افزایش وضوح تصویر و کاهش تاری برای پیش‌پردازش و حذف اشیای اضافی استفاده شده است [۱۸-۱۵]. در محیط‌های دریایی، SSL به‌عنوان مرز بین آسمان و اقیانوس عمل می‌کند و به شناسایی نواحی اقیانوسی و کاهش محدوده جستجو کمک می‌کند [۱۹]. در تشخیص کشتی‌های دور دست، این اهداف معمولاً در نزدیکی خط SSL ظاهر می‌شوند؛ بنابراین، شناسایی این خط پس از مرحله‌ی پیش‌پردازش، فرآیندی مفید و مؤثر به‌شمار می‌آید. روش‌های تشخیص SSL را می‌توان به سه دسته اصلی تقسیم کرد: روش‌های مبتنی بر تبدیل، ویژگی‌های ناحیه‌ای، و بخش‌بندی معنایی [۱۹]. در روش‌های مبتنی بر تبدیل، ابتدا تمامی لبه‌های تصویر استخراج می‌شوند؛ سپس با استفاده از تبدیلات غیرخطی مانند تبدیل هاف، تمامی خطوط موجود در تصویر شناسایی و در نهایت SSL مشخص می‌شود [۲۰]. این روش بر اساس تفاوت‌های موجود در ویژگی‌های تصویر، نواحی مختلف را از یکدیگر تفکیک می‌کند. در روش‌های مبتنی بر ویژگی‌های ناحیه‌ای، مناطق مختلف تصویر بر اساس تفاوت‌های موجود در برخی ویژگی‌ها شناسایی می‌شوند [۲۱]. در روش‌های بخش‌بندی معنایی، پیکسل‌های تصویر برچسب‌گذاری می‌شوند و تقسیم‌بندی دقیق‌تری از نواحی تصویر برای بهبود تشخیص SSL انجام می‌گیرد [۲۲]. گام بعدی در روش‌های مبتنی بر ویژگی، استخراج ROI است. پژوهشگران برای این منظور از روش‌های مختلفی مانند استفاده از اطلاعات SSL، روش‌های آستانه‌گذاری، تشخیص نقاط برجسته، و روش‌های حوزه فرکانس استفاده کرده‌اند [۱۹]. پس از شناسایی ROI، ویژگی‌های خاصی مانند اندازه، شکل، رنگ، بافت و الگوهای موجود در تصویر استخراج می‌شوند [۲۳]. مرحله‌ی نهایی در روش‌های مبتنی بر ویژگی، استفاده از طبقه‌بندهای مختلف برای طبقه‌بندی کشتی‌ها است. از جمله این طبقه‌بندها می‌توان به ماشین بردار پشتیبان [۲۴] و درخت تصمیم‌گیری [۲۵] اشاره کرد.

جدول (۱): خلاصه‌ای از روش‌های طبقه‌بندی کشتی‌ها به همراه مزایا و معایب آنها.

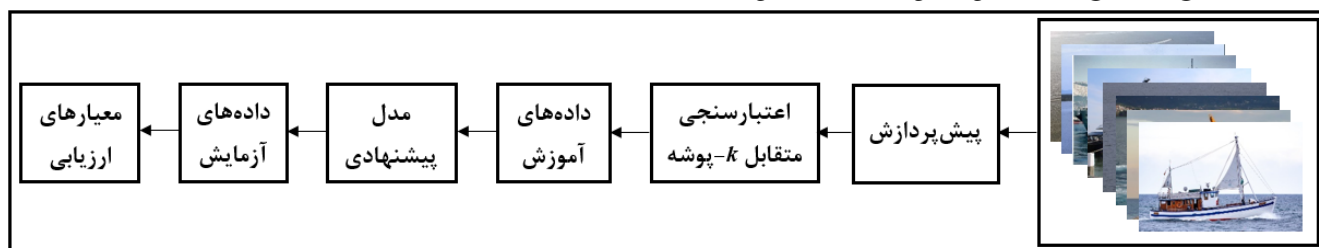
روش	مراحل / انواع	الگوریتم	مزایا	معایب
مبتنی بر ویژگی	پیش‌پردازش	Super-Resolution Techniques[16], Deblur[17, 18]	سهولت اجرا، بدون نیاز به داده‌های برچسب‌دار	نیاز به مهارت و تخصص کافی در انتخاب و استخراج ویژگی، ناتوانی در تعمیم در سناریوهای پیچیده
	تشخیص SSL	Transformation[20], Region Attributes[21], Semantic Segmentation[22]		
	استخراج ROI	SSL Based[27], Threshold Based[19], Saliency Detection[28], Frequency Domain[29]		
	طبقه‌بندی	Slope Entropy[30], Complete Ensemble Empirical[31]		
مبتنی بر یادگیری عمیق	یادگیری نظارت‌شده	CNN[32], Alex Net[33]	یادگیری ویژگی خودکار،	نیاز به مقدار زیادی داده‌های برچسب‌دار، زمان طولانی مرحله آموزش
	یادگیری بدون نظارت	Autoencoders[34]	انتخاب خودکار ویژگی، قابلیت	
	یادگیری انتقالی	GoogleNet[35]	تعمیم در سناریوهای پیچیده	

### ۳- روش پیشنهادی

ارزیابی عملکرد مدل و کاهش سوگیری ناشی از نحوه تقسیم‌بندی داده‌ها می‌شود.

شکل ۱ نمایانگر چارچوب روش پیشنهادی جهت طبقه‌بندی کشتی‌های دریایی است. در گام نخست، تصاویر موجود در مجموعه داده‌ی MARVEL در هشت طبقه‌ی ناوهای هواپیمابر، باربرها، کشتی‌های جنگی، کشتی‌های کروز، کشتی‌های تندرو، کشتی‌های پلیس، کشتی‌های ماهی‌گیری و کشتی‌های لایروبی انتخاب شدند. سپس، این تصاویر تحت فرآیند پیش‌پردازش قرار گرفتند تا کیفیت و کنتراست آنها بهبود یابد. در ادامه، تصاویر پیش‌پردازش شده به مرحله‌ی اعتبارسنجی متقابل  $k$ -پوشه‌ای وارد شدند تا داده‌ها به صورت تصادفی به مجموعه‌های آموزش و آزمون تقسیم شوند. در این پژوهش، از اعتبارسنجی متقابل ۵-پوشه‌ای به منظور ارزیابی عملکرد مدل پیشنهادی استفاده شد. داده‌های آموزشی جهت آموزش مدل مورد استفاده قرار گرفتند و در نهایت، عملکرد مدل بر روی داده‌های آزمون ارزیابی شد. هم‌چنین، ارزیابی نهایی مدل با استفاده از شاخص‌هایی نظیر صحت، دقت و حساسیت انجام پذیرفت.

در مسائل طبقه‌بندی، انتخاب روش مناسب برای ارزیابی عملکرد مدل‌ها اهمیت زیادی دارد. یکی از روش‌های رایج در این زمینه، اعتبارسنجی متقابل  $k$ -پوشه‌ای است که با کاهش وابستگی به یک تقسیم‌بندی خاص از داده‌ها، تخمین دقیق‌تری از عملکرد مدل ارائه می‌دهد [۳۶]. این روش نه تنها باعث کاهش اثرات ناشی از توزیع تصادفی داده‌ها می‌شود، بلکه امکان استفاده حداکثری از کل مجموعه داده را فراهم می‌سازد. علاوه بر این، اعتبارسنجی متقابل موجب می‌شود نتایج ارزیابی از پایداری و دقت بیشتری برخوردار باشند و عملکرد مدل پیشنهادی به صورت منصفانه‌تر با سایر مدل‌ها مقایسه گردد. در این روش، داده‌ها به  $k$  بخش مساوی تقسیم می‌شوند و مدل در  $k$  مرحله‌ی مجزا آموزش می‌بیند و ارزیابی می‌شود. به گونه‌ای که در هر تکرار، یکی از بخش‌ها به عنوان داده‌ی آزمایش و سایر بخش‌های باقی‌مانده برای آموزش مورد استفاده قرار می‌گیرند. با این کار مدل  $k$  بار آموزش می‌بیند و ارزیابی می‌شود، به طوری که تمام نمونه‌ها حداقل یک‌بار در نقش داده‌ی آزمایشی قرار می‌گیرند. این روش باعث افزایش دقت



شکل (۱): چارچوب روش پیشنهادی برای طبقه‌بندی تصاویر کشتی‌های دریایی

باقی بمانند. در پایان مرحله پیش‌پردازش نیز تمامی تصاویر به ابعاد ثابت  $256 \times 256$  پیکسل مقیاس بندی شدند.

### ۱-۳- پیش‌پردازش

در این مرحله، کنتراست تصاویر مورد بررسی قرار گرفت. پس از بررسی مجموعه داده، تصاویری که کنتراست پایین‌تری داشتند با استفاده از روش کنتراست محلی بردارهای فیشر (LCFV) بهبود داده شدند [۳۷]. در این روش، به جای استفاده از کنتراست ساده شدت پیکسل‌ها - که معمولاً فقط تفاوت روشنایی بین نواحی را بررسی می‌کند- از بردارهای فیشر<sup>۲</sup> برای توصیف دقیق‌تر ویژگی‌های سوپر پیکسل‌ها استفاده می‌شود. این ویژگی‌ها با استفاده از مدل آمیخته گاوسی<sup>۳</sup> (GMM) از هر سوپر پیکسل استخراج می‌شوند و شامل اطلاعات آماری چندمرتبه‌ای (یعنی فراتر از میانگین و واریانس ساده) هستند. سپس کنتراست بین این بردارهای فیشر محلی در نواحی مختلف تصویر تحلیل می‌شود تا امکان شناسایی تفاوت‌های جزئی اما مهم بین کشتی و پس‌زمینه‌ی دریا فراهم شود. این روش موجب می‌شود حتی در شرایطی که کشتی‌ها در نوپس زمینه محو می‌شوند، هم‌چنان قابل تفکیک

### ۲-۳- مدل پیشنهادی

مدل پیشنهادی این پژوهش همان‌گونه که در شکل ۲ نشان داده شده است، از چندین لایه پردازشی متوالی تشکیل شده است که با هدف استخراج ویژگی‌های مؤثر و طبقه‌بندی دقیق کشتی‌های دریایی طراحی شده‌اند. این مدل بر پایه ترکیب دو معماری قدرتمند، یعنی شبکه‌های CNN و CapsNet توسعه یافته است. شبکه‌های CNN، که از پرکاربردترین انواع شبکه‌های عمیق در حوزه بینایی ماشین محسوب می‌شوند، با بهره‌گیری از سازوکار لایه‌های کانولوشنی قادر به استخراج ویژگی‌های تصویر در سطوح مختلف هستند. این شبکه‌ها با استفاده از لایه‌های کانولوشن برای شناسایی الگوها، لایه‌های ادغام<sup>۴</sup> برای کاهش ابعاد و لایه‌های تماماً متصل برای تصمیم‌گیری نهایی، عملکرد مؤثری در طبقه‌بندی تصاویر از خود نشان می‌دهند [۳۸].

<sup>۱</sup> Local Contrast of Fisher Vectors

<sup>۲</sup> Fisher Vectors

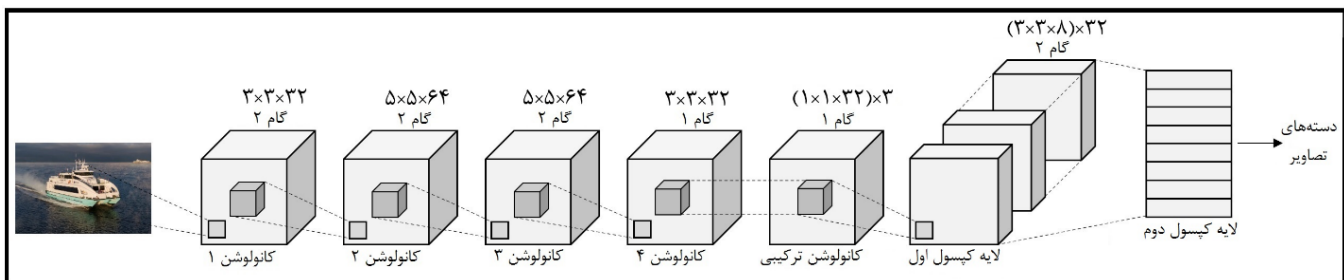
<sup>۳</sup> Gaussian Mixture Model

<sup>۴</sup> Pooling

در تصویر است؛ روابطی که در شرایط چالش‌برانگیز مانند هم‌پوشانی اجسام یا تغییر زاویه دید می‌توانند نقش تعیین‌کننده‌ای در تشخیص صحیح کلاس داشته باشند. این ویژگی موجب شده است که مدل پیشنهادی در این پژوهش که مبتنی بر معماری شبکه‌های CapsNet توسعه یافته است، عملکرد بهتری نسبت به مدل‌های مبتنی بر CNN در طبقه‌بندی کشتی‌های دریایی داشته باشد.

مطابق شکل ۲ مدل پیشنهادی شامل چهار لایه کانولوشنی با تابع فعال‌سازی ReLU، یک لایه کانولوشنی ترکیبی و دو لایه کپسولی است. تصاویر کشتی‌های دریایی با ابعاد  $256 \times 256$  به شبکه وارد می‌شوند و شبکه در طی مرحله آموزش می‌آموزد که آنها را به بردارهای ۳۲ بعدی (۳۲D) کدگذاری کند. جدول ۲ جزئیات لایه‌های مدل پیشنهادی را نشان می‌دهد. طبق این جدول، لایه کانولوشنی اول دارای ۳۲ فیلتر با ابعاد  $3 \times 3$  است که با گام ۲ به تصویر ورودی اعمال می‌شوند. هدف این لایه استخراج ویژگی‌های سطح پایین (مثل لبه‌ها و گوشه‌ها) از تصویر ورودی است. این مرحله برای کاهش ابعاد تصویر و در عین حال حفظ اطلاعات ضروری انجام می‌شود.

شبکه‌های CapsNet نوعی معماری جدید در یادگیری عمیق هستند که با هدف غلبه بر برخی محدودیت‌های ذاتی شبکه‌های CNN طراحی شده‌اند. در حالی که CNNها عملکرد چشم‌گیری در استخراج ویژگی‌ها از تصاویر دارند، اما در بازنمایی روابط فضایی بین اجزای تصویر ناتوان‌اند. به‌عنوان مثال، یک مدل CNN ممکن است وجود یک شیء را تشخیص دهد اما نسبت فضایی بین اجزای آن (مانند موقعیت چشم‌ها نسبت به دهان در تصاویر چهره) را به‌درستی درک نکند. در مقابل، شبکه‌های CapsNet از ساختارهایی به نام کپسول استفاده می‌کنند که هر کدام برداری از ویژگی‌ها را تولید می‌نمایند. این بردار نه تنها نشان می‌دهد که یک ویژگی خاص در تصویر وجود دارد، بلکه اطلاعاتی درباره جهت، موقعیت و مقیاس آن نیز ارائه می‌دهد. مکانیزم مسیریابی پویا بین کپسول‌ها، جایگزین عملگر ادغام در CNN شده است و امکان می‌دهد کپسول‌های پایین‌دستی به‌صورت تطبیقی به کپسول‌های بالادستی متصل شوند. این مکانیزم کمک می‌کند تا ساختارهای سلسله‌مراتبی با دقت بیشتری بازنمایی شوند. مزیت کلیدی شبکه‌های CapsNet در کاربردهایی مانند طبقه‌بندی اهداف دریایی، حفظ روابط فضایی میان اجزای کشتی



شکل (۲): مدل پیشنهادی برای طبقه‌بندی تصاویر کشتی‌های دریایی

نقشه‌های ویژگی با ابعاد  $64 \times 64$  باشد، ادغام سراسری برای هر نقشه ویژگی به صورت رابطه (۱) است:

$$V_{gk} = \sum_{p=0}^{M1} \sum_{q=0}^{M2} n_{kpq} \quad (1)$$

که در آن  $n_k$  تعداد عناصر نقشه‌های ویژگی است. اگر  $V_g$  خروجی لایه ادغام سراسری با ابعاد  $1 \times 1 \times N$  باشد، خروجی‌های لایه‌های کانولوشن با توابع فعال‌سازی Leaky ReLU و Tanh به ترتیب از روابط (۲) و (۳) محاسبه می‌شوند:

$$V_{LR} = \text{LeakyReLU}(\omega_1 \cdot V_g + b_1) \quad (2)$$

$$V_T = \text{Tanh}(\omega_2 \cdot V_g + b_2) \quad (3)$$

به طوری که  $\omega_1$  و  $\omega_2$  ماتریس‌های وزنی و  $b_1$  و  $b_2$  ضرایب ثابت هستند. اگر  $V_{C4}$  خروجی لایه کانولوشن چهارم باشد، خروجی لایه ترکیبی از رابطه (۵) بدست می‌آید

$$V_{HC} = V_{C4} + (V_H * V_{LR}) \quad (4)$$

$$V_H = V_{C4} * V_T \quad (5)$$

پس از لایه اول، سه لایه کانولوشنی دیگر به ترتیب به ترتیب  $64 \times 64$  و  $32 \times 32$  فیلتر با ابعاد  $5 \times 5$  و  $3 \times 3$  وجود دارند. این لایه‌ها ویژگی‌های سطح بالاتر مانند اشکال پیچیده‌تر و ساختارهای متمایز کشتی‌ها را استخراج می‌کنند. هرچه تعداد لایه‌های کانولوشنی بیشتر شود، مدل قادر به یادگیری ویژگی‌های انتزاعی‌تری می‌شود که برای تمایز دقیق‌تر بین کلاس‌ها ضروری است. مقدار گام فیلترها در لایه‌های دوم و سوم نیز ۲ است. اما لایه کانولوشنی چهارم دارای ۳۲ فیلتر با گام ۱ است. هم‌چنین تابع فعال‌سازی در هر چهار لایه کانولوشنی، تابع ReLU است.

لایه پنجم که لایه کانولوشنی ترکیبی نام دارد، از یک لایه ادغام سراسری، یک لایه کانولوشنی با تابع فعال‌سازی Leaky ReLU، یک لایه کانولوشنی با تابع فعال‌سازی Tanh و یک بخش ترکیب‌کننده تشکیل شده است. شکل ۳ جزئیات این لایه را نشان می‌دهد. در این لایه، ادغام سراسری بر روی ۳۲ نقشه ویژگی<sup>۱</sup> ورودی با ابعاد  $64 \times 64$  اعمال می‌شود. اگر ورودی این لایه

داخلی کپسول، هر بردار ورودی در ماتریس وزنی اختصاصی خود با ابعاد  $۸ \times ۳۲$  ضرب می‌شود. این عملیات، فضای ورودی ۸ بعدی کپسول را به فضای خروجی  $۳۲$  بعدی تبدیل می‌کند. خروجی لایه کپسولی ثانویه شامل ۸ بردار  $۳۲$  بعدی است. در طول فرآیند آموزش، برای هر نمونه آموزشی یک مقدار هزینه برای هر یک از این ۸ بردار مطابق با تابع هزینه معادله (۶) محاسبه می‌شود. این تابع هزینه نهایی که با  $L_c$  نشان داده می‌شود، عبارت است از [۱۴]:

$$L_c = T_c \cdot \max(0, m^+ - \|V_c\|)^2 + \lambda(1 - T_c) \cdot \max(0, \|V_c\| - m^-)^2 \quad (۶)$$

به طوری که  $T_c = 1$  نشان‌دهنده برچسب صحیح کلاس  $c$  است.  $m^+$  و  $m^-$  پارامترهای ثابت هستند و  $V_c$  خروجی لایه نهایی است. برای محاسبه هزینه نهایی هر نمونه، هشت مقدار محاسبه شده با هم جمع می‌شوند. بر اساس مدل یادگیری پیشنهادی، هر نمونه در مجموعه داده آموزشی دارای یک برچسب صحیح است که در اینجا به صورت یک بردار ۸ بعدی با ۷ عنصر صفر و ۱ عنصر با مقدار یک (در موقعیت صحیح) ارائه می‌شود. این روش به عنوان کدگذاری One-Hot شناخته می‌شود [۳۹]. در فرمول  $L_c$  (معادله (۶))، برچسب صحیح مقدار  $T_c$  را تعیین می‌کند؛ این مقدار برابر با ۱ خواهد بود اگر برچسب صحیح مربوط به کپسول کلاس خاصی باشد، و در غیر این صورت مقدار آن ۰ خواهد بود. در لایه آخر (لایه کپسول کلاس)، به ترتیب کپسول‌های اول تا هشتم مسئول کدگذاری ناوهای هواپیمابر، باربرها، کشتی‌های جنگی، کشتی‌های کروز، کشتی‌های تندرو، کشتی‌های پلیس، کشتی‌های ماهیگیری و کشتی‌های لایروبی هستند. اگر فرض کنیم که برچسب صحیح مربوط به کلاس ناوهای هواپیمابر باشد، این بدان معناست که کپسول اول در لایه آخر مسئول کدگذاری این کلاس خواهد بود. برای این کپسول کلاس، مقدار  $T_c = 1$  است و برای هفت کپسول باقی‌مانده در این لایه، مقدار  $T_c = 0$  خواهد بود. وقتی  $T_c = 1$  باشد، بخش اول تابع هزینه محاسبه می‌شود و بخش دوم آن صفر است. در مدل پیشنهادی،  $m^- = 0.1$  و  $m^+ = 0.9$  تنظیم شده است.

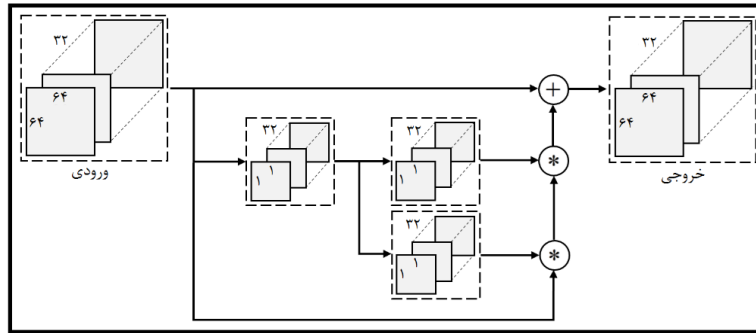
که \* علامت کانولوشن است. لایه ترکیبی با هدف استخراج ویژگی‌های چندمقیاسی از تصاویر طراحی شده است. این لایه شامل چندین مسیر موازی است که هر مسیر فیلترهایی با اندازه‌های مختلف دارد. هر مسیر ویژگی‌های خاصی از تصویر را استخراج می‌کند، به طوری که فیلترهای کوچک‌تر جزئیات دقیق و فیلترهای بزرگ‌تر الگوهای کلی‌تر را شناسایی می‌کنند. سپس ویژگی‌های استخراج شده از طریق عملگرهای جمع و ضرب باهم ترکیب می‌شوند تا یک بازنمایی جامع و غنی از تصویر ایجاد شود. این فرآیند به مدل کمک می‌کند تا هم جزئیات محلی و هم ساختارهای کلی را به طور هم‌زمان شناسایی کند و عملکرد بهتری در طبقه‌بندی تصاویر ارائه دهد.

ویژگی‌های استخراج شده توسط لایه‌های کانولوشنی وارد اولین لایه کپسول می‌شوند. لایه کپسول از واحدهایی به نام کپسول تشکیل شده است که نه تنها حضور یک ویژگی خاص را شناسایی می‌کنند، بلکه موقعیت، مقیاس و جهت آن را نیز مشخص می‌سازند. این خصوصیت باعث می‌شود که مدل در برابر چرخش‌ها، تغییر مقیاس و تغییرات هندسی مقاوم باشد. پس از پردازش اولیه توسط لایه کپسول اول، ویژگی‌های سطح بالاتر به لایه کپسول دوم منتقل می‌شوند. این لایه نقش مهمی در ترکیب ویژگی‌های سطح بالا و ایجاد نمایش برداری برای هر شیء ایفا می‌کند. این نمایش برداری، اطلاعات جامعی از نوع و موقعیت کشتی ارائه می‌دهد.

لایه کپسولی اول شامل  $۳۲$  کپسول است که ویژگی‌های پایه را از لایه‌های کانولوشنی قبلی جمع‌آوری کرده و با هم ترکیب می‌کند. هر کپسول در این لایه از ۸ فیلتر با ابعاد  $۳ \times ۳$  و گام ۲ تشکیل می‌شود. خروجی لایه کپسولی اولیه به صورت ماتریس‌هایی با ابعاد  $(۳۲ \times ۳۲) \times ۸ \times ۳۲$  ارائه می‌شود. لایه کپسولی دوم که به عنوان لایه کپسول کلاس‌ها نیز شناخته می‌شود، شامل ۸ کپسول است که هر کپسول نماینده یکی از کلاس‌های تصاویر کشتی‌های دریایی است. در این لایه، هر کپسول یک نقشه ویژگی  $۳۲$  بعدی  $(۳۲ \times ۳۲) \times ۸ \times ۳۲$  ماتریس را به عنوان ورودی دریافت می‌کند. در واقع، ورودی هر کپسول در این لایه شامل تعداد  $۳۲ \times (۳۲ \times ۳۲) = ۳۲۷۶۸$  بردار ۸ بعدی است. بر اساس عملکرد

جدول (۲): جزئیات لایه‌ها در مدل پیشنهادی

خروجی لایه	تعداد فیلتر/کپسول	ورودی لایه	لایه
نگاشت ویژگی با ابعاد $۲۶۳ \times ۲۶۳$	۳۲ فیلتر با ابعاد $۳ \times ۳$ با گام ۲ و تابع فعالساز ReLU	تصویر با ابعاد $۲۵۶ \times ۲۵۶$	کانولوشن ۱
نگاشت ویژگی با ابعاد $۱۳۱ \times ۱۳۱$	۶۴ فیلتر با ابعاد $۵ \times ۵$ با گام ۲ و تابع فعالساز ReLU	نگاشت ویژگی با ابعاد $۲۶۳ \times ۲۶۳$	کانولوشن ۲
نگاشت ویژگی با ابعاد $۶۵ \times ۶۵$	۶۴ فیلتر با ابعاد $۵ \times ۵$ با گام ۲ و تابع فعالساز ReLU	نگاشت ویژگی با ابعاد $۱۳۱ \times ۱۳۱$	کانولوشن ۳
نگاشت ویژگی با ابعاد $۶۴ \times ۶۴$	۳۲ فیلتر با ابعاد $۳ \times ۳$ با گام ۱ و تابع فعالساز ReLU	نگاشت ویژگی با ابعاد $۶۵ \times ۶۵$	کانولوشن ۴
نگاشت ویژگی با ابعاد $۶۴ \times ۶۴$	ادغام سراسری ۳۲ فیلتر با ابعاد $۱ \times ۱$ با گام ۱ و تابع فعالساز Leaky ReLU ۳۲ فیلتر با ابعاد $۱ \times ۱$ با گام ۱ و تابع فعالساز Tanh	نگاشت ویژگی با ابعاد $۶۴ \times ۶۴$	کانولوشن ترکیبی
ماتریس با ابعاد $۳۲ \times ۳۲ \times ۸ \times ۳۲$	۳۲ کپسول حاوی ۸ فیلتر با ابعاد $۱ \times ۱$ با گام ۱	نگاشت ویژگی با ابعاد $۶۴ \times ۶۴$	کپسول ۱
ماتریس $۸ \times ۳۲$	۸ کپسول ۳۲ بعدی (۳۲D)	ماتریس با ابعاد $۳۲ \times ۳۲ \times ۸ \times ۳۲$	کپسول ۲



شکل (۳): لایه کانولوشن ترکیبی در مدل پیشنهادی

برای هر کلاس ۳۰۰۰ تصویر در نظر گرفته شد. نمونه‌هایی از این ۸ کلاس در جدول ۳ نشان داده است. نمونه‌هایی از تصاویر با کنتراست بهبود یافته نیز در جدول ۴ نشان داده شده است. همان‌طور که از این جدول مشخص است، بهبود کنتراست باعث می‌شود که مرز بین کشتی و دریا و همچنین مرز بین کشتی با سایر اشیاء واضح‌تر شود.

#### ۲-۴- معیارهای ارزیابی

برای نشان دادن عملکرد مدل پیشنهادی از معیارهای کمی مختلفی مانند صحت<sup>۲</sup>، دقت<sup>۳</sup>، حساسیت<sup>۴</sup>، ویژگی<sup>۵</sup> و امتیاز F1 استفاده شد. صحت نشان می‌دهد مدل تا چه اندازه خروجی را به درستی پیش‌بینی کرده است. دقت سطح اطمینان مدل را در خروجی پیش‌بینی شده نمایش می‌دهد. حساسیت عملکرد طبقه‌بند را با در نظر گرفتن تعداد رخدادها یک کلاس خاص نشان می‌دهد. ویژگی معیاری است که نرخ صحیح طبقه‌بندی شده کلاس‌های مخالف کشتی را نشان می‌دهد. امتیاز F1 میانگین هارمونیک دقت و حساسیت است که تعادلی بین این دو معیار برای ارزیابی عملکرد مدل طبقه‌بندی فراهم می‌کند. بر اساس ماتریس درهم‌آمیختگی، نمایش ریاضی معیارهای ارزیابی انتخاب شده بیان می‌شود [۴۱]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (9)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (10)$$

$$\text{F1-score} = \frac{2TP}{2TP + FP + FN} \quad (11)$$

Accuracy<sup>۲</sup>  
Precision<sup>۳</sup>  
Sensitivity<sup>۴</sup>  
Specificity<sup>۵</sup>

#### ۴- نتایج

در این بخش نتایج بدست آمده از مدل پیشنهادی بررسی می‌شود. مقاله حاضر به طبقه‌بندی تصاویر کشتی‌های دریایی با استفاده از شبکه‌های کپسولی پرداخته است. به همین منظور ابتدا مجموعه داده مورد استفاده معرفی می‌شود. در ادامه معیارهای کمی مختلفی مثل دقت، صحت و حساسیت برای نشان دادن عملکرد مدل پیشنهادی در طبقه‌بندی تصاویر کشتی‌های دریایی معرفی می‌شوند. همچنین ماتریس درهم‌آمیختگی به منظور ارزیابی عملکرد طبقه‌بندی تصاویر ارائه می‌گردد. در ادامه نمونه‌هایی از تصاویر با پیش‌بینی درست و غلط نیز ذکر می‌شود. در نهایت مقایسه مدل پیشنهادی با سایر مدل‌های موجود در مطالعات گذشته انجام شد. قابل ذکر است که تمامی آزمایش‌ها بر روی سیستمی با پردازنده ۶۴ بیتی Intel Core i7 و ۸ گیگابایت رم با سیستم عامل ویندوز ۱۱ انجام شد.

#### ۱-۴- مجموعه داده

در طبقه‌بندی کشتی‌های دریایی به‌ویژه هنگام استفاده از چارچوب‌های یادگیری عمیق، در اختیار داشتن یک مجموعه داده‌ی مناسب و قابل آموزش، نخستین و ضروری‌ترین گام به‌شمار می‌رود. مجموعه داده‌های تصویری بزرگ معمولاً تعداد کافی تصاویر را برای آموزش یک مدل عمیق پیچیده فراهم می‌کنند. با این حال تهیه تصاویر و یا ویدیوهای با کیفیت بالا به دلیل محیط‌های پیچیده دریایی دشوار است. به همین دلیل در این مقاله از مجموعه داده MARVEL استفاده شد [۴۰]. مجموعه داده MARVEL شامل ۲ میلیون تصویر کشتی دریایی است که برای هر تصویر ویژگی‌هایی مانند عرض، پرچم، شماره سازمان بین‌المللی دریانوردی، نام، طول، کلاس، ظرفیت حمل تانکرها و نوع کشتی ثبت شده است. در این مجموعه داده ۱۰۳۷۰۱ شماره سازمان بین‌المللی دریانوردی<sup>۱</sup> (IMO) یکتا وجود دارد [۴۰]. از بین این مجموعه داده ۸ کلاس ناوهای هواپیمابر، باربرها، کشتی‌های جنگی، کشتی‌های کروز، کشتی‌های تندرو، کشتی‌های پلیس، کشتی‌های ماهیگیری و کشتی‌های لایروبی انتخاب شدند و

تشخیص داده است. FN بیانگر تعداد تصاویری از کلاس هدف است که مدل آن‌ها را نادرست در کلاس‌های دیگر طبقه‌بندی کرده است. در نهایت، TN شامل تصاویری از سایر کلاس‌هاست که مدل به درستی آن‌ها را در کلاس‌های واقعی خود (غیر از کلاس هدف) قرار داده است. در این مقاله، عملکرد مدل پیشنهادی بر پایه این شاخص‌ها از نظر پایداری، اثربخشی و توانایی تعمیم‌پذیری مورد ارزیابی قرار گرفته است.

که به ترتیب نمایانگر صحت، دقت، حساسیت، ویژگی و امتیاز F1 هستند. با توجه به ماهیت چندکلاسه‌ی مسئله طبقه‌بندی کشتی‌ها، مقادیر TP، TN، FP و FN برای هر کلاس به صورت مستقل محاسبه شدند. در این چارچوب، TP نشان‌دهنده‌ی تعداد تصاویری از کلاس هدف است که مدل به درستی آن‌ها را در همان کلاس طبقه‌بندی کرده است. FP تعداد تصاویری از سایر کلاس‌هاست که مدل به اشتباه آن‌ها را متعلق به کلاس هدف

جدول (۳): نمونه‌هایی از تصاویر منتخب از مجموعه داده MARVEL

تصاویر				کلاس
				ناوهای هواپیمابر
				باربرها
				کشتی‌های جنگی
				کشتی‌های کروز
				کشتی‌های تندرو
				کشتی‌های پلیس
				کشتی‌های ماهیگیری
				کشتی‌های لایروبی

جدول (۴): نمونه‌هایی از تصاویر پیش‌پردازش شده

				تصویر اصلی
				تصویر پیش‌پردازش شده

جدول (۵): مقادیر انتخابی برای هایپر پارامترهای اصلی

مقدار انتخابی	پارامتر
۰/۰۰۱	نرخ یادگیری
۳۲	اندازه هر دسته
۵۰	تعداد دوره
Adam	بهینه‌ساز
۸	ابعاد بردار خروجی هر کپسول
۳	تکرار الگوریتم مسیریابی پویا

ماتریس درهم‌آمیختگی ابزاری برای ارزیابی عملکرد مدل‌های طبقه‌بندی است که در آن ردیف‌ها نشان‌دهنده کلاس‌های واقعی (برچسب‌های صحیح) و ستون‌ها نشان‌دهنده کلاس‌های پیش‌بینی شده توسط مدل هستند. هر عنصر این ماتریس تعداد نمونه‌هایی را نشان می‌دهد که مدل آنها را از کلاس واقعی به کلاس پیش‌بینی شده تخصیص داده است. عناصر روی قطر اصلی ماتریس بیانگر تعداد نمونه‌هایی هستند که مدل به درستی آنها را طبقه‌بندی کرده است، در حالی که عناصر خارج از قطر، نمونه‌هایی هستند که به اشتباه در کلاس‌های دیگر طبقه‌بندی شده‌اند. شدت رنگ سلول‌های این ماتریس نشان‌دهنده تعداد نمونه‌ها در هر کلاس است، به طوری که رنگ‌های تیره‌تر نشان‌دهنده تعداد بیشتر نمونه‌ها می‌باشد. این ماتریس به شناسایی الگوهای خطا، کلاس‌هایی که بیشترین اشتباه را دارند و عملکرد کلی مدل در طبقه‌بندی کلاس‌های مختلف کمک می‌کند.

ماتریس درهم‌آمیختگی برای مدل پیشنهادی در شکل ۴ نشان داده شده است. همان‌طور که از این شکل استنتاج می‌شود، الگوریتم پیشنهادی توانسته است در ۹۷/۶۱٪ موارد کشتی‌های دریایی را به درستی طبقه‌بندی کند. بر اساس این شکل از ۹۰۰ تصویر موجود برای کلاس ناوهای هواپیمابر، ۸۵۹ تصویر بدرستی طبقه‌بندی شدند. اما از ۴۱ تصویر باقیمانده ۳۹ تصویر به اشتباه در کلاس باربرها طبقه‌بندی شدند. این می‌تواند بخاطر زاویه برخی از تصاویر ناوهای هواپیمابر باشد که از زاویه کناری شبیه کشتی‌های باربر

### ۳-۴ - طبقه‌بندی کشتی‌های دریایی

همان‌طور که پیش‌تر ذکر شد در این مقاله برای ارزیابی عملکرد مدل پیشنهادی، از روش اعتبارسنجی متقابل ۵-پوشه‌ای استفاده شد. در این روش، کل مجموعه داده (۲۴۰۰۰ تصویر) به پنج بخش مساوی تقسیم گردید. در هر تکرار، ۸۰٪ داده‌ها (چهار بخش یا ۱۹۲۰۰ تصویر) برای آموزش مدل و ۲۰٪ داده‌ها (یک بخش یا ۴۸۰۰ تصویر) به عنوان داده‌های اعتبارسنجی مورد استفاده قرار گرفت. این فرآیند به گونه‌ای انجام شد که هر بخش دقیقاً یک‌بار به عنوان داده‌های اعتبارسنجی انتخاب و مدل روی تمامی تصاویر آزمایش شود. استفاده از اعتبارسنجی متقابل موجب می‌شود تا عملکرد مدل به طور جامع‌تری ارزیابی گردد. در هر تکرار، معیارهای ارزیابی از جمله دقت، صحت، حساسیت محاسبه شدند و در نهایت، میانگین این معیارها به عنوان عملکرد نهایی مدل گزارش گردید.

برای آموزش مدل پیشنهادی، مجموعه‌ای از هایپر پارامترها انتخاب و تنظیم شدند که در عملکرد نهایی مدل نقش مهمی ایفا می‌کنند. به منظور تعیین دقیق تنظیمات مدل، مقادیر انتخابی برای هایپر پارامترهای اصلی در جدول ۵ درج شده است. نرخ یادگیری برابر با ۰/۰۰۱ در نظر گرفته شده که مقدار مناسبی برای همگرایی پایدار در شبکه‌های عمیق است. تعداد نمونه‌های هر دسته برابر با ۳۲ تنظیم شد تا تعادل مناسبی بین مصرف حافظه و دقت آموزش برقرار شود. هم‌چنین، مدل به مدت ۵۰ دوره آموزش داده شد که با بررسی رفتار تابع هزینه، همگرایی مناسبی حاصل شد. برای به‌روزرسانی وزن‌ها از الگوریتم بهینه‌سازی Adam استفاده شد که در شبکه‌های عمیق عملکرد مؤثری دارد. در بخش شبکه‌های کپسولی، ابعاد بردار خروجی هر کپسول برابر با ۸ و تعداد تکرار الگوریتم مسیریابی پویا برابر با ۳ در نظر گرفته شد که با نتایج مقالات پیشین و بررسی تجربی تطابق دارد.

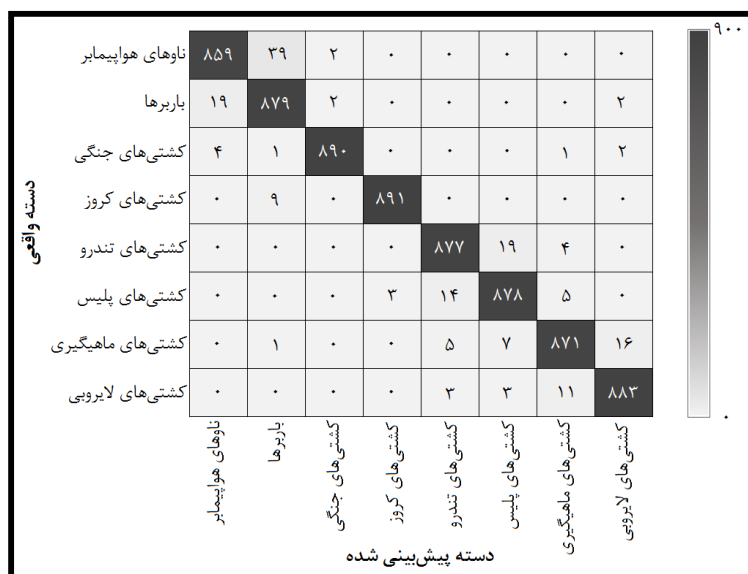
Batch Size <sup>۱</sup>Epoch <sup>۲</sup>

شناخته شده مانند AlexNet، GoogleNet، ResNet-50 و VGGNet مقایسه شده است. ابتدا از نظر درصد صحت کلی مشاهده می‌شود که مدل پیشنهادی با صحت ۹۷/۶۱٪ عملکردی رقابتی و قابل توجه ارائه می‌دهد و نسبت به بیشتر مدل‌های پایه موجود بهتر عمل می‌کند. به‌ویژه نسبت به AlexNet که صحت کمتری دارد و حتی GoogleNet که معمولاً عملکرد بالایی دارد، مدل پیشنهادی توانسته است بهبود مناسبی در درصد صحت نشان دهد. از نظر حساسیت، که بیانگر توانایی مدل در شناسایی صحیح تصاویر متعلق به کلاس هدف است، مدل پیشنهادی با ثبت مقدار ۹۶/۸۰٪ عملکرد بسیار بالایی داشته است. این مقدار نشان می‌دهد که مدل حتی در شرایط پیچیده و چالش‌برانگیز نیز قادر به شناسایی دقیق کشتی‌های مورد نظر بوده است.

زمان پاسخ‌دهی مدل پیشنهادی ۳۵ میلی‌ثانیه است که با توجه به ساختار پیچیده‌تر شبکه کپسولی، زمان مناسبی محسوب می‌شود. این زمان در مقایسه با مدل‌های VGGNet و ResNet که زمان‌های پاسخ‌دهی بالاتری دارند، کارایی بهتری را در کاربردهای بلادرنگ نشان می‌دهد. از منظر زمان آموزش، مدل پیشنهادی با حدود ۶ ساعت، نسبت به مدل‌های پیچیده‌تر مانند ResNet-50 (۱۴ ساعت) و VGGNet (۲۵ ساعت) زمان کمتری نیاز دارد و این مزیت در پروژه‌های عملی و توسعه سریع‌تر مدل بسیار مهم است. البته مدل AlexNet که ساختار ساده‌تری دارد، زمان آموزش کوتاه‌تری دارد اما صحت و حساسیت پایین‌تر آن، کارایی کلی مدل را محدود می‌کند. بنابراین، مدل پیشنهادی تعادلی مطلوب بین صحت بالا، حساسیت قوی، زمان آموزش نسبتاً کوتاه و زمان پاسخ‌دهی قابل قبول برقرار کرده است. این ویژگی‌ها، مدل را به گزینه‌ای مناسب برای سیستم‌های نظارتی هوشمند در محیط‌های دریایی تبدیل می‌کند که نیازمند پردازش سریع و دقیق تصاویر در شرایط واقعی هستند.

می‌شود و مدل را دچار اشتباه می‌کند. همچنین در کلاس باربرها نیز تعداد ۱۹ تصویر آن به اشتباه در کلاس ناوهای هواپیمابر طبقه‌بندی شده است. تصاویر کشتی‌های تندرو و پلیس نیز از بعضی زوایا شبیه به هم هستند و مدل را در تشخیص درست سردرگم خواهند کرد. این وضعیت برای کشتی‌های ماهیگیری و لایروبی نیز تکرار شده است. بالاترین دقت طبقه‌بندی مربوط به کلاس کشتی‌های کروز و پایین‌ترین آن مربوط به کلاس کشتی‌های ناوهای هواپیمابر به ترتیب با صحت‌های ۹۹٪ (۸۹۱ تشخیص درست) و ۹۵/۴۴٪ (۸۵۹ تشخیص درست) است. جدول ۶ نمونه تصاویر کشتی‌های دریایی را با پیش‌بینی درست و غلط توسط مدل پیشنهادی نشان می‌دهد. همان‌طور که از این شکل پیداست تصاویر کشتی‌های پلیس با تصاویر کشتی‌های تندرو و یا حتی در زاویه خاص دوربین با تصاویر کشتی‌های ماهیگیری می‌تواند اشتباه گرفته شود. البته این اشتباه می‌تواند به دلیل شرایط جوی عکس نیز رخ دهد. به طور مثال در ستون آخر جدول ۶ تصویر کشتی لایروبی توسط مدل به عنوان کشتی جنگی شناخته شده است که این به دلیل شرایط جوی و عدم وضوح مناسب تصویر گرفته شده است.

نتایج طبقه‌بندی توسط مدل‌های شناخته‌شده در جدول ۷ ثبت شده است. این جدول مقایسه‌ای بین مدل‌های مختلف طبقه‌بندی در سال‌های اخیر ارائه داده است. اگر چه تعداد کلاس‌ها و تعداد تصاویر استفاده شده در مدل‌های مختلف یکسان نیست، اما مدل پیشنهادی به دقت بالاتری نسبت به سایر مدل‌ها دست یافته است. مدل‌هایی که دارای تعداد تصاویر کمتری هستند، از یادگیری انتقالی استفاده کرده‌اند. در واقع تعداد تصاویر ذکر شده برای آنها تعداد تصاویر برای آزمایش بوده است و مدل عمیق آنها روی مجموعه داده‌های بزرگتری آموزش دیده است. در جدول ۸، عملکرد مدل پیشنهادی با چند شبکه عمیق پایه و



شکل (۳): ماتریس درهم‌ریختگی برای طبقه‌بندی توسط مدل پیشنهادی

جدول (۶): نمونه‌هایی از تصاویر کشتی با پیش‌بینی صحیح و ناصحیح توسط مدل پیشنهادی

کشتی‌های لایروبی	کشتی‌های ماهیگیری	کشتی‌های تندرو	ناوهای هواپیمابر	
				درست
 پیش‌بینی: کشتی‌های جنگی	 پیش‌بینی: کشتی‌های پلیس	 پیش‌بینی: کشتی‌های پلیس	 پیش‌بینی: کشتی‌های لایروبی	غلط

جدول (۷): مقایسه مدل‌های مختلف در طبقه‌بندی مجموعه داده MARVEL.

مدل	تعداد داده	تعداد گروه‌ها	سال	دقت کلی
EfficientNet-B5 [12]	~۱۱۷k	۲۶	۱۳۹۹	۹۱/۳۹
Modified Resnet 152 [13]	~۱k	۵	۱۴۰۰	۹۱/۳۵
Cas-Ship Net [11]	~۹k	۵	۱۴۰۰	۹۵/۰۶
Optimized CNN [10]	~۳k	۵	۱۴۰۲	۹۷/۰۴
مدل پیشنهادی	~۲۴k	۸	۱۴۰۳	۹۷/۶۱

جدول (۸): مقایسه زمانی مدل‌های شناخته‌شده با مدل پیشنهادی

مدل	صحت (%)	حساسیت (%)	زمان پاسخ‌دهی مدل (میلی ثانیه)	زمان آموزش (ساعت)
GoogleNet	۹۶/۵۰	۹۵/۵۰	۳۰	۹
AlexNet	۹۵/۲۰	۹۴/۰۰	۲۵	۴
ResNet	۹۷/۰۰	۹۶/۲۰	۴۰	۱۴
VGGNet	۹۶/۸۰	۹۵/۹۰	۴۵	۲۵
مدل پیشنهادی	۹۷/۶۱	۹۶/۶	۳۵	۵

## ۵- بحث

GoogleNet در برخی کلاس‌ها مانند کلاس کشتی‌های کروز صحت خوبی به دست آورده است اما همچنان نسبت به مدل پیشنهادی صحت کمتری دارد. علاوه بر این، تعمیم‌پذیری مدل پیشنهادی تحت شرایط غیر استاندارد و جوی نامناسب مانند حالت شب، مه، وجود اشیای متعدد در تصویر یا فاصله زیاد هدف تا دوربین ارزیابی شد. نمونه‌ای از این تصاویر برای هر کلاس در جدول ۱۰ ارائه شده است. برای این منظور، ۵۰ تصویر با شرایط غیر استاندارد از هر کلاس در مجموعه داده MARVEL انتخاب شدند. این تصاویر به مدل آموزش‌دیده ارائه شدند و نتایج در جدول ۱۱ ثبت شد. قابل ذکر است که این تصاویر در فرآیند آموزش مدل وجود نداشتند. همان‌طور که در این جدول مشاهده می‌شود، مدل پیشنهادی در تمامی کلاس‌ها حداقل ۴۰ نمونه را به درستی طبقه‌بندی کرده است که نشان‌دهنده عملکرد مناسب و قابل قبول آن است.

برای ارزیابی دقیق‌تر عملکرد مدل پیشنهادی، از چند شبکه‌ی عمیق شناخته‌شده و پایه جهت مقایسه استفاده شد. برای این منظور شبکه‌های GoogleNet، AlexNet، ResNet-50 و VGGNet ابتدا با استفاده از مجموعه داده منتخب که شامل ۳۰۰۰ تصویر در هر کلاس بود، آموزش داده شدند. سپس شبکه‌های آموزش‌دیده برای پیش‌بینی ۸ کلاس انتخابی در این مقاله مورد استفاده قرار گرفتند. در هر کلاس، ۹۰۰ تصویر آزمایشی مشابه (۳۰٪ از مجموعه داده قبلی) برای این شبکه‌ها استفاده شد. درصد صحت برای هر کلاس در جدول ۹ ثبت شده است. همان‌طور که از این جدول مشاهده می‌شود، شبکه پیشنهادی در تمامی کلاس‌ها نسبت به سایر شبکه‌های شناخته‌شده درصد صحت بالاتری دارد. البته به دلیل پیچیدگی و تعداد لایه‌های زیاد، این شبکه‌ها اغلب به زمان قابل‌توجهی برای آموزش نیاز دارند. با این وجود، شبکه

جدول (۹): درصد صحت تصاویر آزمایش برای هر کلاس در مجموعه داده MARVEL

پیشنهادی	VGG Net	ResNet	Alex Net	Google Net	مدل کلاس
۹۵/۴۴	۹۰/۴۴	۸۹/۸۸	۸۷/۶۶	۹۲/۱۱	ناوهای هواپیما
۹۷/۶۶	۹۲/۲۲	۹۲/۸۸	۸۹/۵۵	۹۴/۳۳	باربرها
۹۸/۸۸	۹۳/۲۲	۹۴/۵۵	۹۱/۸۸	۹۴/۴۴	کشتی‌های جنگی
۹۹/۰۰	۹۴/۳۳	۹۴/۳۳	۹۱/۳۳	۹۴/۵۵	کشتی‌های کروز
۹۷/۴۴	۹۲/۶۶	۹۲/۲۲	۸۹/۸۸	۹۱/۸۸	کشتی‌های تندرو
۹۷/۵۵	۹۱/۸۸	۹۱/۸۸	۹۰/۸۸	۹۲/۲۲	کشتی‌های پلیس
۹۶/۷۷	۹۱/۰۰	۹۰/۵۵	۸۹/۳۳	۹۱/۵۵	کشتی‌های ماهیگیری
۹۸/۱۱	۹۲/۷۷	۹۱/۴۴	۹۱/۱۱	۹۲/۰۰	کشتی‌های لایروبی

جدول (۱۰): نمونه تصاویر با شرایط خاص (حالت شب، مه، اشیاء متعدد، فاصله طولانی)

کلاس	نمونه تصاویر	کلاس	نمونه تصاویر	کلاس
ناوهای هواپیما		باربرها		ناوهای هواپیما
کشتی‌های جنگی		کشتی‌های کروز		کشتی‌های جنگی
کشتی‌های تندرو		کشتی‌های پلیس		کشتی‌های تندرو
کشتی‌های ماهیگیری		کشتی‌های لایروبی		کشتی‌های ماهیگیری

جدول (۱۱): تعداد تصاویر با پیش‌بینی صحیح و ناصحیح تحت شرایط خاص

کلاس	ناوهای هواپیما	باربرها	کشتی‌های جنگی	کشتی‌های کروز	کشتی‌های تندرو	کشتی‌های پلیس	کشتی‌های ماهیگیری	کشتی‌های لایروبی
طبقه‌بندی درست	۴۱	۴۰	۴۴	۴۶	۴۲	۴۱	۴۱	۴۲
غلط	۹	۱۰	۶	۴	۸	۹	۹	۸

کلاس‌های اقلیت وزن بیشتری دریافت می‌کنند. در نتیجه مدل در فرآیند آموزش، خطاهای مربوط به کلاس‌های اقلیت را جدی‌تر تلقی کرده و تلاش می‌کند آنها را بهتر شناسایی کند. در پیاده‌سازی این رویکرد، وزن‌های کلاس‌ها براساس نسبت معکوس تعداد نمونه‌ها محاسبه و به تابع خطای مورد استفاده در مدل اعمال شد. این مکانیزم کمک می‌کند تا مدل از گرایش بیش‌ازحد به کلاس‌های اکثریت جلوگیری کند و تعادل بهتری در فرآیند یادگیری حاصل شود.

جدول ۱۲ مقاومت مدل پیشنهادی را در برابر عدم‌توازن داده‌ها بررسی می‌کند. طبق این جدول کاهش تعداد نمونه‌های کلاس اقلیت باعث افت قابل‌توجه در عملکرد مدل شده است؛ به‌ویژه معیارهایی مانند حساسیت و امتیاز F1 بیشترین کاهش را داشته‌اند. به‌عنوان مثال، در حالت کاهش به ۱۰٪ و بدون تابع وزن‌دار، حساسیت این کلاس به حدود ۶۲٪ کاهش یافته که نشان از عدم توانایی مدل در شناسایی صحیح این کلاس اقلیت دارد. همچنین، صحت کلی مدل نیز به‌صورت تدریجی کاهش یافته است که به‌دلیل کاهش عملکرد کلی شبکه در مواجهه با کلاس‌های نامتوازن است. با این حال، اعمال تابع هزینه وزن‌دار در تمام نسبت‌های نامتوازنی باعث بهبود محسوس در شناسایی کلاس اقلیت شده است. به‌طور خاص، امتیاز F1 کلاس اقلیت در حالت ۱۰٪ از ۶۵۸/۰ به ۷۶۲/۰ افزایش یافته است که نشان‌دهنده تأثیر مثبت وزن‌دهی بر فرآیند یادگیری مدل است. این بهبود بدون افت چشمگیر در صحت کلی حاصل شده و نشان می‌دهد که مدل می‌تواند با بهره‌گیری از تکنیک‌های مناسب، در شرایط سخت‌تری نیز عملکرد قابل‌قبولی داشته باشد. در مجموع، این آزمایش بیانگر آن است که مدل پیشنهادی نه‌تنها در شرایط متوازن عملکرد مطلوبی دارد، بلکه با استفاده از روش‌هایی مانند تابع هزینه وزن‌دار، می‌تواند تا حد زیادی در برابر چالش عدم‌توازن داده‌ها مقاوم باشد و طبقه‌بندی قابل‌اعتمادی ارائه دهد.

در بسیاری از مسائل دنیای واقعی، مجموعه‌داده‌ها از نظر تعداد نمونه در کلاس‌های مختلف به‌طور یکنواخت توزیع نشده‌اند [۴۲]. این پدیده که با عنوان طبقه‌بندی نامتوازن شناخته می‌شود، زمانی رخ می‌دهد که یک یا چند کلاس دارای تعداد نمونه به‌مراتب کمتر از سایر کلاس‌ها باشند [۴۳]. در چنین شرایطی الگوریتم‌های یادگیری که معمولاً بر مبنای بیشینه‌سازی صحت کلی آموزش داده می‌شوند، تمایل دارند کلاس‌های غالب (کلاس‌های اکثریت) را بیشتر یاد بگیرند؛ در نتیجه عملکرد ضعیفی در شناسایی کلاس‌های کم‌نمونه (کلاس‌های اقلیت) نشان می‌دهند. این عدم تعادل می‌تواند منجر به کاهش صحت، یادگیری ناقص ویژگی‌های کلاس‌های اقلیت، و در نهایت بی‌اعتباری مدل در کاربردهای مهم مثل شناسایی اسپرم‌های نابارور و یا تشخیص ترک در سازه‌های بتنی شود [۴۴-۴۵]. از این‌رو، طراحی روش‌هایی برای مقابله با این چالش مانند استفاده از تکنیک‌های بازنمونه‌گیری و یا تابع هزینه وزن‌دار ضروری و رایج است.

به‌منظور تحلیل تأثیر عدم توازن در عملکرد مدل پیشنهادی، یکی از کلاس‌ها به صورت کنترل‌شده به‌عنوان کلاس اقلیت در نظر گرفته شد. برای این منظور، تعداد نمونه‌های متعلق به کلاس کشتی‌های کروز به‌طور عمدی کاهش یافت تا یک سناریوی نامتوازن شبیه‌سازی شود. تعداد نمونه‌های این کلاس به ترتیب به ۲۵، ۵۰ و ۱۰٪ مقدار اصلی کاهش یافت. سپس مدل در این وضعیت‌ها آموزش داده شد و عملکرد آن با استفاده از تابع هزینه وزن‌دار مورد ارزیابی قرار گرفت. این آزمایش امکان بررسی پایداری مدل در شرایط عدم توازن و سنجش اثربخشی مکانیزم جبرانی پیشنهادی را فراهم کرد.

در شرایطی که داده‌ها به‌صورت نامتوازن توزیع شده‌اند، استفاده از تابع هزینه وزن‌دار روشی مؤثر برای بهبود عملکرد مدل محسوب می‌شود. در این روش، به هر کلاس ضربی متناسب با معکوس فراوانی آن در مجموعه‌داده نسبت داده می‌شود، به‌طوری‌که

جدول (۱۲): بررسی معیارها در حالت طبقه‌بندی نامتوازن

نسبت نمونه کلاس اقلیت	تابع وزن‌دار	صحت کلی (%)	F1 کلاس اقلیت	F1 میانگین	حساسیت کلاس اقلیت
۱۰۰٪ (متوازن)	×	۹۷/۶۱	۰/۹۷۶	۰/۹۷۰	۰/۹۶۸
۵۰٪ از مقدار اولیه	×	۹۵/۴۰	۰/۹۵۳	۰/۸۹۲	۰/۸۷۵
۵۰٪ از مقدار اولیه	✓	۹۶/۱۰	۰/۹۶۱	۰/۹۲۵	۰/۹۱۸
۲۵٪ از مقدار اولیه	×	۹۳/۰۵	۰/۹۲۸	۰/۷۹۳	۰/۷۶۵
۲۵٪ از مقدار اولیه	✓	۹۴/۵۰	۰/۹۴۵	۰/۸۵۶	۰/۸۴۰
۱۰٪ از مقدار اولیه	×	۹۰/۲۰	۰/۹۰۱	۰/۶۵۸	۰/۶۲۰
۱۰٪ از مقدار اولیه	✓	۹۲/۳۰	۰/۹۲۳	۰/۷۶۲	۰/۷۴۵

## ۶- نتیجه‌گیری

در دهه اخیر، با پیشرفت صنعت دریایی، به‌کارگیری فناوری‌های هوش مصنوعی در کاربردهای مختلف دریایی از جمله وسایل سطحی خودکار، کشتی‌های خودران و پایش هوشمند ترافیک دریایی ضروری شده است. یکی از جنبه‌های کلیدی نظارت دریایی، شناسایی و طبقه‌بندی انواع کشتی‌هاست که به دلیل تنوع ساختاری و ویژگی‌های بصری آنها، چالشی پیچیده محسوب می‌شود. در این پژوهش، یک مدل هوشمند مبتنی بر شبکه‌های کپسولی با بهره‌گیری از لایه ترکیبی جهت استخراج ویژگی‌های چندمقیاسی از تصاویر ارائه شد که توانایی حفظ روابط فضایی بین ویژگی‌ها را دارد و به درک بهتر ساختار کلی اشیاء منجر می‌شود.

تحلیل ماتریس درهم‌آمیختگی نشان داد که مدل پیشنهادی عملکرد مطلوبی در طبقه‌بندی کشتی‌ها دارد؛ به طوری که بالاترین صحت مربوط به کشتی‌های کروز با مقدار ۹۹٪ و پایین‌ترین صحت برای ناوهای هواپیمابر با ۹۵/۴۴٪ بود که همچنان نسبت به مدل‌های پایه، نتیجه قابل قبولی محسوب می‌شود. همچنین مقایسه مدل پیشنهادی با شبکه‌های عمیق شناخته‌شده مانند GoogleNet، AlexNet، ResNet و VGGNet بیانگر برتری مدل در صحت کلی و عملکرد در تمامی کلاس‌ها بود.

ارزیابی مدل در شرایط غیر استاندارد (شامل شب، مه، حضور اشیای متعدد و فواصل دور) نشان داد که مدل علاوه بر صحت بالا، قابلیت تعمیم‌پذیری مناسبی نیز دارد. در بررسی داده‌های نامتوازن، رویکرد کاهش عمده نمونه‌های یکی از کلاس‌ها و استفاده از تابع هزینه وزن‌دار، اثر مثبت بر حفظ حساسیت و امتیاز F1 کلاس اقلیت داشت و نشان داد که مدیریت نامتوازی داده‌ها برای افزایش قابلیت اعتماد مدل اهمیت بالایی دارد.

با توجه به محدودیت‌های موجود در داده‌های برجسب‌دار و چالش‌های محیطی، بهره‌گیری از یادگیری خودنظارتی می‌تواند به استخراج ویژگی‌های عمیق‌تر و مقاوم‌تر کمک کند و مدل را در مواجهه با داده‌های کم‌برجسب و نامتوازن توانمند سازد. همچنین به‌کارگیری رویکردهای ترکیبی شامل شبکه‌های کپسولی همراه با روش‌های افزایش داده می‌تواند موجب بهبود قابلیت تعمیم‌پذیری سیستم طبقه‌بندی در شرایط مختلف شود. این رویکردها چشم‌اندازهای جدیدی برای ارتقای عملکرد سامانه‌های طبقه‌بندی کشتی‌ها در محیط‌های دریایی پیچیده فراهم می‌کنند.

## مراجع

- images: A literature survey," *Remote sensing of environment*, vol. 207, pp. 1–26, 2018.
- [3] R. Danovaro *et al.*, "Ecological variables for developing a global deep-ocean monitoring and conservation strategy," *Nature Ecology & Evolution*, vol. 4, no. 2, pp. 181–192, 2020.
- [4] L. Bo, X. Xiaoyang, W. Xingxing, and T. Wenting, "Ship detection and classification from optical remote sensing images: A survey," *Chinese Journal of Aeronautics*, vol. 34, no. 3, pp. 145–163, 2021.
- [5] X. Chen, "AI and big data: Leveraging machine learning for advanced data analytics," *Advances in Computer Sciences*, vol. 7, no. 1, 2024.
- [6] T. Yang, J. Chen, and N. Zhang, "AI-empowered maritime Internet of Things: A parallel-network-driven approach," *IEEE Network*, vol. 34, no. 5, pp. 54–59, 2020.
- [7] A. Rawson and M. Brito, "A survey of the opportunities and challenges of supervised machine learning in maritime risk analysis," *Transport Reviews*, vol. 43, no. 1, pp. 108–130, 2023.
- [8] W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K.-R. Müller, "Explaining deep neural networks and beyond: A review of methods and applications," *Proceedings of the IEEE*, vol. 109, no. 3, pp. 247–278, 2021.
- [9] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 12, pp. 6999–7019, 2021.
- [10] M. H. Salem, Y. Li, Z. Liu, and A. M. AbdelTawab, "A transfer learning and optimized CNN based maritime vessel classification system," *Applied Sciences*, vol. 13, no. 3, p. 1912, 2023.
- [11] F. Ucar and D. Korkmaz, "A novel ship classification network with cascade deep features for line-of-sight sea data," *Machine Vision and Applications*, vol. 32, no. 3, p. 73, 2021.
- [12] M. H. Salem, Y. Li, and Z. Liu, "Transfer learning on efficientnet for maritime visible image classification," in *2022 7th international conference on signal and image processing (ICSIP)*, 2022: IEEE, pp. 514–520.
- [13] L. A. Leonidas and Y. Jie, "Ship classification based on improved convolutional neural network architecture for intelligent transport systems," *Information*, vol. 12, no. 8, p. 302, 2021.
- [14] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," *Advances in neural information processing systems*, vol. 30, 2017.
- [15] E. Teixeira, B. Araujo, V. Costa, S. Mafra, and F. Figueiredo, "Literature review on ship localization, classification, and detection methods based on optical
- [1] R. Lou, Z. Lv, S. Dang, T. Su, and X. Li, "Application of machine learning in ocean data," *Multimedia systems*, vol. 29, no. 3, pp. 1815–1824, 2023.
- [2] U. Kanjir, H. Greidanus, and K. Oštir, "Vessel detection and classification from spaceborne optical

- the motion attitude of visible light sensors," *Sensors*, vol. 19, no. 18, p. 4004, 2019.
- [28] C. Lin, W. Chen, and H. Zhou, "Multi-visual feature saliency detection for sea-surface targets through improved sea-sky-line detection," *Journal of Marine Science and Engineering*, vol. 8, no. 10, p. 799, 2020.
- [29] A. Zhou, W. Xie, and J. Pei, "Infrared maritime target detection using the high order statistic filtering in fractional Fourier domain," *Infrared Physics & Technology*, vol. 91, pp. 123-136, 2018.
- [30] Y. Li, P. Gao, B. Tang, Y. Yi, and J. Zhang, "Double feature extraction method of ship-radiated noise signal based on slope entropy and permutation entropy," *Entropy*, vol. 24, no. 1, p. 22, 2021.
- [31] H. Yang, L.-l. Li, G.-h. Li, and Q.-r. Guan, "A novel feature extraction method for ship-radiated noise," *Defence Technology*, vol. 18, no. 4, pp. 604-617, 2022.
- [32] X. Xu, X. Zhang, and T. Zhang, "Multi-scale SAR ship classification with convolutional neural network," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, pp. 4284-4287, 2021.
- [33] T. Lu, B. Han, and F. Yu, "Detection and classification of marine mammal sounds using AlexNet with transfer learning," *Ecological Informatics*, vol. 62, p. 101277, 2021.
- [34] M. Liang, R. W. Liu, S. Li, Z. Xiao, X. Liu, and F. Lu, "An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation ", *Ocean Engineering*, vol. 225, p. 108803, 2021.
- [35] X. Du, Y. Sun, Y. Song, H. Sun, and L. Yang, "A Comparative Study of Different CNN Models and Transfer Learning Effect for Underwater Object Classification in Side-Scan Sonar Images," *Remote Sensing*, vol. 15, no. 3, p. 593, 2023.
- [36] L. A. Yates, Z. Aandahl, S. A. Richards, and B. W. Brook, "Cross validation for model selection: a review with examples from ecology," *Ecological Monographs*, vol. 93, no. 1, p. e1557, 2023.
- [37] X. Wang, G. Li, X.-P. Zhang ,and Y. He, "Ship detection in SAR images via local contrast of Fisher vectors," *IEEE Transactions on geoscience and remote sensing*, vol. 58, no. 9, pp. 6467-6479, 2020.
- [38] S. Pawan and J. Rajan, "Capsule networks for image classification: A review," *Neurocomputing*, vol. 509, pp. 102-120, 2022.
- [39] Z. Lv, H. Ding, L. Wang, and Q. Zou, "A convolutional neural network using dinucleotide one-hot encoder for identifying DNA N6-methyladenine sites in the rice genome," *Neurocomputing*, vol. 422, pp. 214-221 ,2021.
- sensors and neural networks," *Sensors*, vol. 22 ,no. 18, p. 6879, 2022.
- [16] Z. Wang, J. Chen, and S. C. Hoi, "Deep learning for image super-resolution: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp. 3365-3387, 2020.
- [17] P. Satish, M. Srikantaswamy ,and N. K. Ramaswamy, "A Comprehensive Review of Blind Deconvolution Techniques for Image Deblurring," *Traitement du Signal*, vol. 37, no. 3, 2020.
- [18] R. Singh and S. Bansal, "A comparative study of image deblurring techniques," *Journal of Computational and Theoretical Nanoscience*, vol. 17, no. 9-10, pp. 4571-4579, 2020.
- [19] L. Wang *et al.*, "A review of methods for ship detection with electro-optical images in marine environments," *Journal of Marine Science and Engineering*, vol. 9, no. 12, p. 1408, 2021.
- [20] J. Fu, J. Zhao, and F. Li, "Infrared sea-sky line detection utilizing self-adaptive Laplacian of Gaussian filter and visual-saliency-based probabilistic Hough transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2021.
- [21] D .Liang and Y. Liang, "Horizon detection from electro-optical sensors under maritime environment," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 1, pp. 45-53, 2019.
- [22] W. Yang, H. Li, J. Liu, S. Xie, and J. Luo, "A sea-sky-line detection method based on Gaussian mixture models and image texture features," *International Journal of Advanced Robotic Systems*, vol. 16, no. 6, p. 1729881419892116, 2019.
- [23] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction," *Journal of Applied Science and Technology Trends*, vol. 1, no. 1, pp. 56-70, 2020.
- [24] S. Wu, X. Chen, C. Shi, J. Fu, Y. Yan, and S. Wang, "Ship detention prediction via feature selection scheme and support vector machine (SVM)," *Maritime Policy & Management*, vol. 49, no. 1, pp. 140-153, 2022.
- [25] L. R. Abreu, I. S. Maciel, J. S. Alves, L. C. Braga, and H. L. Pontes, "A decision tree model for the prediction of the stay time of ships in Brazilian ports," *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105634, 2023.
- [26] F. Zhuang *et al.* , "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43-76, 2020.
- [27] X. Shan, D. Zhao, M. Pan, D. Wang, and L. Zhao, "Sea-Sky line and its nearby ships detection based on



**حامد جبجاری** تحصیلات خود را در رشته مهندسی برق-کنترل در مقطع کارشناسی در دانشگاه صنعتی شاهرود و در مقطع کارشناسی ارشد و دکتری در دانشگاه بین‌المللی امام خمینی(ره) قزوین به پایان رساند. ایشان در سال ۱۴۰۱ در مقطع دکتری فارغ‌التحصیل شد. زمینه‌های پژوهشی و تخصصی وی پردازش تصویر، داده‌کاوی و هوش مصنوعی است.



**حمید هوشمند** مدرک کارشناسی خود را در رشته مهندسی رباتیک از دانشگاه صنعتی شاهرود در سال ۱۳۸۷ و مدرک کارشناسی ارشد خود را در رشته مهندسی مکترونیک از دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران در سال ۱۳۹۰ دریافت کرد. وی فارغ‌التحصیل مقطع دکتری از دانشگاه صنعتی شاهرود است. علائق پژوهشی و تخصص ایشان رباتیک و هوش مصنوعی، کنترل غیرخطی، کاهش مرتبه مدل و کنترل سیستم‌های مکترونیک است.



**نوشین بیگدلی** کارشناسی و کارشناسی ارشد مهندسی برق خود را در دانشگاه شیراز به ترتیب در گرایش‌های الکترونیک و کنترل به پایان رساند و در سال ۱۳۸۶ در مقطع دکتری برق-کنترل از دانشگاه صنعتی شریف فارغ‌التحصیل گردید. ایشان هم‌اکنون، استاد گروه مهندسی برق-کنترل دانشگاه بین‌المللی امام خمینی (ره) قزوین است. زمینه‌های تخصصی وی عبارتست از: داده‌کاوی، استخراج ویژگی‌ها و آنالیز سری زمانی، پردازش تصاویر و علائم، سیستم‌های هوشمند، کنترل سیستم‌های آشوبگونه و کنترل پیش‌بین مدل.

[40] E. Gundogdu, B. Solmaz, V. Yücesoy, and A. Koc, "Marvel: A large-scale image dataset for maritime vessels," in *Computer Vision-ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part V 13*, 2017: Springer, pp. 165-180.

[41] A. Ali-Gombe and E. Elyan, "MFC-GAN: Class-imbalanced dataset classification using multiple fake class generative adversarial network," *Neurocomputing*, vol. 361, pp. 212-221, 2019.

[42] H. Jabbari and N. Bigdeli, "New conditional generative adversarial capsule network for imbalanced classification of human sperm head images," *Neural Computing and Applications*, vol. 35, no. 27, pp. 19919-19934, 2023.

[43] H. Jabbari and N. Bigdeli, "A new hierarchical algorithm based on CapsGAN for imbalanced image classification," *IET Image Processing*, vol. 18, no. 1, pp. 194-210, 2024.

[44] H. Jabbari and N. Bigdeli, "A New Capsule Generative Adversarial Network for Imbalanced Classification of Human Sperm Images," *Journal of Modeling in Engineering*, vol. 21, no. 73, pp. 279-294, 2023.

[۴۵] جبجاری، حامد و بیگدلی، نوشین، "طراحی و ارزیابی یک شبکه عصبی کپسولی جدید برای طبقه‌بندی نامتوازن تصاویر"، *مجله ماشین بینایی و پردازش تصویر*، دوره ۹، شماره ۱، صفحه ۱۵-۱، فروردین ۱۴۰۱.